Categoricals:

* Borough
* Community District (CD) – need to transform somehow
* Census Tract ( CT2010 ~1300 uniques)
* Census Block (CB2010 ~350 uniques)
* School District (SchoolDist)
* City Council District (Council)
* ZipCode (~200 uniques)
* Fire Company (FireComp – categorize based on type E, L, Q, etc)
* Police Precinct (PolicePrct)
* Health Area (HealthArea ~200 uniques)
* Sanitation District Boro (SanitBoro)
* Sanitation District Number (SanitDist)
* Sanitation Subsection (SanitSub)
* Address
* Zoning District 1 (ZoneDist1) – can use for residential/commercial
* Zoning District 2, Zoning District 3, Zoning District 4, Overlay1, Overlay2, SPDist1, SPDist2, LtdHeight - Too Many NaN
* All Zoning Components 1 (AllZoning1)
* AllZoning2 - Too Many NaN
* Split Binary Indicator (Binary)
* Building Class
* Land Use – indicates residential vs. commercial
* Owner Type, Owner Name, Area Source – Too many NaN and don’t think we need
* Extension Code (Ext)
* Proximity Code (ProxCode)
* Irregular Code (IrrLotCode - binary)
* Lot Type (LotType)
* Basement Type (BsmtCode)
* Year Built Code (BuiltCode)
* Historic District Name (HistDist) – Many NaN’s
* Landmark (Landmark) – Many NaN’s
* BoroCode (repetitive)
* Borough, Tax Block and Lot (BBL) – **USE TO MERGE**
* Condominium Number (CondoNo)
* Census Tract 2 ( Tract2010 ~1300 uniques)
* Zoning Map # (ZoneMap)
* Zoning Map Code (ZMCode) – Too many NaN’s
* Sanborn Map # (Sanborn)
* Tax Map # (TaxMap)
* E-Designation # (EDesigNum) – Too many NaN’s
* Apportionment BBL (APPBBL)
* Apportionment Date (convert to continuous or get rid of)
* Pluto Base Map Indicator (PLUTOMapID), Version Number (Version) – Don’t think we need

Questions:

* What to do with features like Garage Area? Add to total floor area?
* Change Year Altered 1 and Year Altered 2 to reflect what we want it to?